(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2001 (26.04.2001)

PCT

(10) International Publication Number
WO 01/29252 A2

(51) International Patent Classification⁷: C12Q 1/68

(21) International Application Number: PCT/EP00/10268

(22) International Filing Date: 18 October 2000 (18.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
99120618.6 18 October 1999 (18.10.1999) EP

(71) Applicant and
(72) Inventor: SPEICHER, Michael [DE/DE]; Waldparkstrasse 70a, 85521 Riemerling (DE).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): EILS, Roland [DE/DE]; Strahlenberger Strasse 26, 69198 Schriesheim (DE).

(74) Agent: VOSSIUS & PARTNER; Siebertstrasse 4, 81675 München (DE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/29252 A2

(54) Title: A METHOD FOR THE GENERATION OF REPEAT-DEPLETED DNA

(57) Abstract: The invention relates to a method for the generation of repeat-depleted DNA comprising amplifying repetitive template DNA by a first polymerase chain reaction (PCR), wherein the hybridization step is a low stringency hybidization step and a second PCR followi ng the first PCR, wherein the hybridization step of said second PCR is a high stringency hybridization step. The repeat-depleted DNA obtained can be used as probe or cloned into vectors, plasmid, etc. Further, the invention relates to the application of the method in the generation of and hybridization with DNA libraries, DNA arrays or DNA blots.

A method for the generation of repeat-depleted DNA

The invention relates to a method for the generation of repeat-depleted DNA comprising amplifying repetitive template DNA by a first polymerase chain reaction (PCR), wherein the hybridization step is a low stringency hybridization step and a second PCR following the first PCR, wherein the hybridization step of said second PCR is a high stringency hybridization step. The repeat-depleted DNA obtained can be used as probe or cloned into vectors, plasmid, etc. Further, the invention relates to the application of the method in the generation of and hybridization with DNA libraries, DNA arrays or DNA blots.

The human genome is composed of 3 billion base pairs organized into 24 physically distinct genetic units called chromosomes. All genes are arranged linearly along the chromosomes, and most human cells contain two sets of chromosomes, one set from each parent. Early characterization of the complexity of genomic DNA by studying reassociation kinetics ($Cot_{1/2}$) led to the understanding that complex genomes contain DNA sequences that are unique in the genome (unique DNA) and DNA that is present as multiple copies in the genome (repetitive DNA). In human genomes, unique DNA ($Cot_{1/2}$ value=1000 moles of nucleotide seconds/liter) makes up half of the total DNA. Among the repetitive DNA is a fraction that reanneals with a Cot value of between 100 and 1000, the moderately repetitive DNA, which includes sequences for the major RNA and protein structural components of the cell, such as genes for rRNA and tRNA (several hundred copies per genome) and genes encoding the histone H1 protein (about 400 per haploid genome) (Szabo, 1992). The most highly repetitive DNA in the human genome is the so-called satellite DNA ($Cot_{1/2}$ value as low as 10^{-3}), which is composed of a very long series of tandem repetitions of a short nucleotide sequence. Satellite DNA is often found in the heterochromatin associated with the centromere chromosomes, where it may contain several thousands homologous repeat units.

Highly repetitive satellite DNA is also interspersed (singly or in tandem arrays) throughout the genome. Historically, the interspersed repeats have been grouped into two classes: the short interspersed nuclear elements (SINES), consisting of repeats shorter than 500 bp, and the long interspersed nuclear elements (LINES). The Alu family of genetic elements is 300 bp SINE, with about 900,000 copies (3 to 6% of the genome) dispersed throughout the genome, giving an average distance between copies of about 4 kb. The Alu sequence is believed to have become dispersed throughout the genome by a translocation mechanism that is partially encoded within the element. The L1 genetic element is a LINE sequence found at about 20,000 to 50,000 per genome, but consists of elements averaging 6 to 7 kb in size.

However, the high amount of repetitive DNA is disadvantageously for wide range of applications.

For the amplification of unique DNA, it is easy to calculate the number of expected targets when designing an amplification experiment using human genomic DNA: 1 µg of genomic DNA contains about 152,000 genome equivalents. For unique DNA, this means 152,000 template molecules, but for the repetitive tRNA genes, about $1.5x10^6$ copies and about 10^{11} copies of the Alu element. Thus, in normal polymerase chain reactions (PCR) with a complex probe as PCR template the repetitive elements represent the vast majority of DNA. Although PCR was initially introduced to amplify the single locus in a target DNA, it is increasingly being used to amplify multiple loci simultaneously. Several PCR methods have been developed to produce even and uniform amplified chromosomal DNA. The interspersed repetitive sequence-PCR (IRS-PCR) has in this manner been used to create human chromosomes and region specific libraries (Nelson, 1989, Lotter, 1990). The use of primers which anneal to repetitive sequences within the template DNA to be amplified allows the amplification of the segments between suitably positioned repeats. However, each amplification product after IRS-PCR contains repetitive sequences. Further, in LA-PCR (linker adapted PCR) fragments resulting from a restriction digest are ligated to short oligonucleotides which serve as primer sites for general PCR across unknown sequences. Telenius, 1992 (Genomics 13, 718-725) describes a version of the PCR, termed "degenerated oligonucleotide primed PCR" (DOP-PCR), which employs

oligonucleotides of partially degenerated sequences. This degeneracy, together with the PCR protocol utilizing a low initial annealing temperature, ensures priming from multiple (e.g., ~10^6 in human) evenly dispared sites within a given genome. According to Telenius, DOP-PCR represent a rapid, efficient and species-independent technique for general DNA amplification. Furthermore, Telenius, 1992 states that DOP PCR produces a more even and uniform amplification than Alu-PCR or IRS-PCR, which are directed towards repeat-rich regions. Therefore, according to Telenius, 1992, DOP-PCR products show a repeated to single copy sequence ratio which represents the chromosomal distribution. Recently, Kunkasjärvi, 1997 (genes, chromosomes and cancer, 18, 94 to 101) published an optimized DOP-PCR for the universal amplification of small DNA samples containing repetitive sequences. However, DNA comprising repetitive sequences has a disadvantage in several amplifications, e.g. in situ hybridization (FISH) or for the preparation of genomic DNA libraries.

In particular, the ability to hybridize complex DNA probes to specific metaphase chromosomes and interphase nuclei has made FISH an indispensable tool for both clinical diagnostics and basic research (for reviews, see Carter 1996, and Lichter 1997). The vast majority of probes used in fluorescence in situ hybridization contain competitive DNA. Traditionally, the specific hybridization of complex probes is achieved by the addition of excess quantities of an unlabeled blocking agent (or competitor DNA), such as genomic DNA, different fractions of repetitive DNA, and Cot-1 DNA (Landegent et al. 1987; Lichter et al. 1988; Pinkel et al. 1988). Commercial Cot-1 DNA is expensive, however. Further, a certain fraction of labeled repetitive sequences well hybridize to target sequences, particularly in the case of suboptimal conditions of cot-1 DNA supression. In cases of incomplete suppression, the signal-to-background ratio is low, making the identification of signals difficult or impossible. Moreover, specialized FISH applications have evolved over the last few years that require very large amounts of competitor DNA. These applications include the hybridization of YAC clones after Alu-PCR (Lengauer et al. 1992); comparative genomic hybridization (CGH) (Kallioniemi et al. 1992; du Manoir et al. 1993), particularly if extended to matrix/microarray-CGH (Solinas-Toldo et al. 1997; Pinkel et al. 1998); and procedures using combinatorially labeled painting probes, such as

multiplex-FISH (M-FISH) (Speicher et al. 1996) and spectral karyotyping (SKY) (Schröck et al. 1996).

Thus, to avoid the use of a blocking agent, different approaches were developed. One approach achieved specific hybridization of a limited number of probes by extended self-preannealing prior to hybridization (Wienberg et al. 1997). A potential new probe source for the painting of human chromosomes without suppression could be chromosome paints derived from hominoid primates (Müller et al. 1997). Here, no blocking agent is needed because during evolution repetitive sequences diverged at a higher rate than single-copy sequences (Warburton and Willard 1996). However, a complete set of these probes to paint all human chromosomes is not yet available.

Recently, protocols based on affinity chromatography and positive (Chen-Liu et al. 1995; Rouquier et al. 1995) and negative (Craig et al. 1997) subtraction hybridization were developed to generate FISH probes depleted of repetitive sequences. The approach described by Craig et al. (1997) consists of three steps. In the first step, biotin-labeled subtractor-DNA is hybridized to a source DNA in solution. Next, subtractor-source hybrids are separated from nonhybridized source DNA molecules using affinity chromatography. Then, the nonhybridized source DNA is purified and amplified by PCR, and the quality of depletion is tested in a FISH experiment. If the desired degree of selection has not yet been achieved, further rounds of depletion can be performed, either with the same subtractor DNA again or with a different DNA source as a subtractor. Further improvements of this technology and the generation of a complete set of human painting probes depleted of repetitive sequences were described by Bolzer et al. (1999). All of these probes yield highly specific signals and can be hybridized without competitor DNA and without preannealing. However, the above mentioned methods are time-consuming, dependent on skilled and trained personnel, and expensive to perform for a large number of probes. Thus, an easy and cheap method to provide improved probes should find widespread applications.

Thus, the technical problem underlying the present invention is to provide means and methods which allow the production of substantially repeat-depleted DNA.

The solution to this technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the present invention relates to a method for the generation of repeat-depleted DNA comprising amplifying repetitive template DNA by a first polymerase chain reaction (PCR), wherein the hybridization step is a low stringency hybridization step, and a second PCR following the first PCR, wherein the hybridization step of said second PCR is a high stringency hybridization step. Surprisingly, it was found, that due to the much faster hybridization kinetics of repetitive sequences repetitive DNA fragments hybridize faster compared to single copy sequences. Under appropriate PCR conditions this effect causes that repetitive elements are already double stranded by the time that the polymerase starts with the synthesis of new nucleotide strands. As a result, single copy DNA is preferentially amplified and the obtained DNA is repeat-depleted. Products of said PCR amplification were tested with Southern-blots after the first PCR. Hybridization with a pan-alphoid probe which stains simultaneously all centromeres in a metaphase spread did surprisingly not show any signals. In contrast, hybridization with a Cot-1 DNA probe yielded signals indicating that repetitive sequences are still present in the probe mix.

The term "repeat-depleted" as used herein refers to a polynucleotide which contains a reduced ratio of repetitive/single copy sequences in comparison to said ratio of the template for its generation. The production of a repeat-depleted DNA might be performed in one step or via one or more intermediates in more than one step. The intermediates can have the same, a reduced or increased percentage of repetitive sequences compared to the probe. Preferably, repeat-depleted DNA comprises about 5% or more less repetitive sequences than the used template, more preferably, the repeat-depleted DNA comprises about 10% or more less, even more preferred is a reduction of about 20% or more, even more preferred are about 30% or more, more preferred are about 50% or more and most preferred is a reduction of about 70% or more. The person skilled in the art knows to analyze the amount of repetitive DNA, for example, by determining the hybridization efficiency of said DNA to highly repetitive sequences, e.g., centromers or by cloning the DNA obtained into vectors and determining the number of clones containing a repetitive sequence (Telenius, 1992 and references therein) A simple test to check the presence of repetitive DNA in a probe is to perform a FISH experiment without preanneling and competitor DNA. A probe free of repetitive elements should result in a specific signal without

background. A more laborious technique was described in the Craig et al. (1996) and Bolzer et al. (1999) papers. After hybridizing a probe with biotinylated Cot-1 DNA in solution the double-stranded repetitive elements can be captured with magnetic beads coupled to avidin. The amount of captured DNA can be measured either on an agarose gel or with a spectrophotometer.

The term "low stringency hybridization step" as used herein refers to conditions used in the annealing step of the PCR. Standard settings of conditions for PCR are well within the skill of the artisan and published, see for example, Sambrook et al. (Molecular Cloning; A Laboratory Manual, second edition, Coldspring Harbor Laboratory Press, Coldspring Harbor, New York (1989) or Erhardt, PCR protocols). Thus, the person skilled in the art knows to determine the optimal hybridization temperature for the used primer according to well-known protocols and common theoretical considerations (Sambrook, 1989). Based on said laboratory manual knowledge and published theoretical considerations for calculating an optimal primer annealing temperature the person skilled in the art knows to test variation of said annealing temperature and to determine an annealing temperature which allows the preferred amplification (see, e.g., Mullis, Birghäuser, PCR, 1994). Said theoretical consideration can be, for example, based on the G/C rule (Sambrook, 1989). The estimation of the annealing-temperature is a standard procedure in molecular genetics. In fact, nowadays most PCR-machines have a Tm-Calculator which performs the calculation. For example, the GeneAmp-PCR System 9700 machine (PE Applied Biosystems) has a calculator that performs above mentioned calculations using information such as salt concentration, primer concentration, and primer sequences.

The term "repetitve template DNA" relates to DNA comprising sequences which occur many times within the genome, e.g., alphoid sequences, satellite-DNA, mini-, microsatellites, etc.

The most highly repetitive DNA in the human genome is the so-called satellite DNA ($Cot_{1/2}$ value as low as 10^{-3}), which is composed of a very long series of tandem repetitions of a short nucleotide sequence. Satellite DNA is often found in the heterochromatin associated with the centromere chromosomes, where it may contain several thousands homologous repeat units.

Highly repetitive satellite DNA is also interspersed (singly or in tandem arrays) throughout the genome. Historically, the interspersed repeats have been grouped into two classes: the short interspersed nuclear elements (SINES), consisting of repeats shorter than 500 bp, and the long interspersed nuclear elements (LINES). The Alu family of genetic elements is 300 bp SINE, with about 900,000 copies (3 to 6% of the genome) dispersed throughout the genome, giving an average distance between copies of about 4 kb. The Alu sequence is believed to have become dispersed throughout the genome by a translocation mechanism that is partially encoded within the element. The L1 genetic element is a LINE sequence found at about 20,000 to 50,000 per genome, but consists of elements averaging 6 to 7 kb in size. A detailed overview about human repetitive elements was recently published by Lee et al. (1997).

Here, the term "low stringency hybridization step" relates to a temperature in the hybridization step, which is around 10°C or more below the optimal annealing temperature of that primer, preferably the temperature is around 15°C or more below, even more preferably are 20°C or more below the optimal annealing temperature. However, the annealing temperature should not be below about 25°C, preferably a low stringency hybridization step occurs at about 30°C.

Under "high stringency hybridization conditions" is understood that the temperature in said step is preferably about 15°C or more below or above the optimal annealing temperature of said primer. More preferably, the hybridization step temperature is about 10°C or more below or above and most preferred is a hybridization step temperature which is of about the optimal annealing temperature. However, "high stringency temperature" according to the present invention relates to a temperature in the annealing step of the second PCR which is at least about 10°C or more above the hybridization conditions of the first PCR, more preferred are about 15°C or more and most preferred is a hybridization temperature which is about 20°C or more above the hybridization temperature of the first PCR.

The term "primer" as used herein refers to an oligonucleotide whether occurring naturally as in a purified restriction digest or produced synthetically. The primer is preferably single-stranded for a maximum of efficiency in the method of the present invention, and is preferably an oligodesoxyribonucleotide. Purification of said primers is generally envisaged, prior to their use in the method of the present invention. Such

purification steps can comprise HPLC (high performal liquid chromatography) or PAGE (polyacrylamid gel electrophoresis), and are known to the person skilled in the art. Oligonucleotides, representing primers as used in method of the present invention, can be identified, obtained and tested according to the state of the art which is visually represented by computer based sequence analysis and laboratory manuals, e.g., Sambrook, 1989.

Primers used in accordance with the present invention anneal to chromosomal DNA or genomic DNA, multiple times and are e.g. capable to prime the amplification of a complete genome under appropriate conditions. Such primers can be degenerated or partially degenerated primers, e.g., 6MW primers [SEQ ID NO: 1]. The method of the present invention can be performed if the primer anneals to frequent sides in the repetitive sequences comprising DNA, e.g., CTA-4PCR primers (Crack et al., 1997) [SEQ ID NO: 2]. About 100 nanograms of products have been amplified by the CTA-4PCR technique. Further primers to be used according to the present invention can also anneal specifically to distinct chromosomes. However, primers which are annealing specifically to the repetitive sequences of the template DNA (e.g. IRS-PCR primers) cannot be used according to the present invention.

The term "hybridized to" in accordance with the present invention denote the pairing of two polynucleotide strands by hydrogen bonding between complementary nucleotides.

In a preferred embodiment the method of the present invention comprises a preannealing step. The "term preannealing" step relates to the hybridization of the probe in solution prior to its application, e.g., on a slide.

An additional preannealing can be used to achieve highly specific results in case that for time limitation a probe has to be generated very quickly and there is not sufficient time for the required rounds of reamplifications and a cross-hybridization is observed at other chromosomes. However, the PCR alone does always suppress alphoid sequences sufficiently after the first round of PCR, thus virtually any DNA probe has a reduced number of repetitive elements. Therefore, a probe which may yield unsatisfactory results yields after a very short preannealing time of five minutes or less already satisfactory results. For example, the hybridization mix can be taken directly from the heat block or placed on ice and then dropped on the hybridization

field very quickly to ensure that absolute no preannealing could happen. This is for many routine labs cumbersome: usually the tubes are removed from the heat block and then the probe mixture is dropped on the slide. Thus, under the conditions used by most operators the tubes are exposed to room temperature for various times (in the range of 1 to several minutes).

In a preferred embodiment the transition time between the hybridization step and the elongation step is between about 2 and about 8 minutes, further preferred are between about 3 and about 6 minutes. In a more preferred embodiment, the transition time is between about 5 and about 6 minutes. In a more preferred embodiment said transition time is about 6 minutes.

In particular, it was found that an extension of the transition time between the hybridization step and the elongation step ("ramp" time) of the first PCR results in a further decreased amplification of repetitive sequences. In accordance with the knowledge of a person skilled in the art the term "hybridization step" relates to the PCR step in which the primer can anneal to the template of the PCR. The term "elongation step" relates to the step of the PCR which occurs at or about a temperature in which the polymerase emzyme used for the amplification of the DNA has its preferred activity, e.g., its highest activity or an activity with a decreased error rate. For example, Taq polymerase has an optimal amplification activity at 72°C. However, other DNA polymerases can also be used, e.g., Sequenase (USP, Cleveland, OH), Taq Plus polymerase (Stratagene, La Jolla, CA), ThermoSequenase, Pwo polymerase, etc. The use of a specific polymerase determines the selection of an optimal extension step temperature and length.

However, in regard of the theoretical considerations described above, the transition time can also be omitted without loosing the capability to obtain repeat-depleted DNA.

According to another embodiment, the first PCR comprises more than 3 amplification cycles, more preferred are more than 4 amplification cycles, even more preferred are between 5 and 10 cycles.

In a further preferred embodiment the hybridization step of the first PCR is between about 1 and 8 minutes, preferably the hybridization step is between about 1.5 minutes and 6 minutes in length. In an even more preferred embodiment, the hybridization step takes place for about 2 minutes.

It is well in the skills of an artisan to determine an optimal hybridization time as well as to determine an optimal transition time between the hybridization step and the elongation step by using standard test systems, as e.g. described in the examples.

Further advantageous conditions for performing the method of the present invention depend on the primer and the template used. For example, as regards the cycle number and the time periods of the steps of the second PCR each step can be varied according to the knowledge of the person skilled in the art and laboratory manual description (Sambrook, 1989, Erlich, H.A., PCR Technology, Principles and Applications for DNA Amplification).

Furthermore, advantageous PCR conditions might depend on the use of the repeat-depleted DNA. For example, in the case the repeat-depleted DNA is used for the generation of a DNA library a specific length of the PCR fragments obtained might be advantageous. Thus, the person skilled in the art will vary according normal laboratory manual knowledge the parameters of the PCR, e.g., use a proof reading polymerase and/or elongate the extension step to obtain longer DNA fragments and/or add, e.g., 1 second or more per cycle to the time period of the extension step and/or change the annealing temperature to achieve more or less stringent conditions. Accordingly, non-limiting examples for the generation of repeat-depleted DNA for the use as FISH probes are shown in the pending examples of the present application.

In another embodiment of the present invention said first and/or second PCR can be repeated once or more times. It was found that repeating the steps of the method of the present invention results advantageously in a further depletion of repetitive sequences. Thus, in a preferred embodiment the first and/or second PCR is/are repeated once, more preferably the first and/or second PCR is/are repeated twice. Even more preferred is repeating of the first and/or second PCR three times. However, the method of the present invention can even be repeated more than three times.

In another preferred embodiment the method of the present invention is for the generation of a repeat-depleted DNA of higher or lower eukaryots, e.g., of mammals (e.g., humans, rats, mice, etc.), plants, fungi or protozoa.

The method of the present invention can be used for the generation of repeat-depleted DNA from many different origins. Studies of the kinetics of the re-association of formerly denatured DNA revealed that eukaryotic DNA, in contrast to prokaryotic DNA, contains many repeated base sequences. Highly repetitive sequences are found in all animals' genomes, e.g., from human to fungi, as well as in higher plants. Until now, the usage of, e.g., hybridization methods as FISH, is mainly restricted to the investigation of that few organisms for which Cot-1 DNA is commercially available. Therefore, it is advantageous to provide a method which overcome the lack of availability of Cot-1 DNA for many organisms. The present invention provides a method for the production of repeat-depleted DNA wherein the DNA of most organisms can be used as template.

In one embodiment the present invention relates to a repetitive template DNA which is e.g. an YAC, cosmid, PAC, BAC, or plasmid. Furthermore, said repetitive template DNA can be isolated from cell lines, e.g., hybrid cell lines, or isolated by microdissecting or generated by other PCR protocols. Furthermore, the repetitive template DNA can be total or partial genomic DNA, or is derived from plastids as, e.g., chloroplast DNA or mitochondria DNA. Specific chromosomes might be isolated, e.g., by microdissection, and then be used as template for the generation of repeat-depleted DNA, e.g., for the generation of chromosome specific painting probes.

Accordingly to another embodiment, the repeat-depleted DNA of the present invention is ligated into plasmids, vectors, yeast artificial chromosomes (YACs), cosmids, P1-derived artificial chromosomes (PACs), or Bacterial artificial chromosomes (BACs).

In another embodiment the repeat-depleted DNA is further immobilized on a suitable support. Accordingly, the present invention relates to the use of the repeat-depleted DNA of the present invention in combination with chip technologies, i.e. for the

preparation of DNA arrays, as well as for DNA blots. Such a support can be a membrane, e.g., a nylon or nitrocellulose membrane or other supports comprising reactive groups, (see, e.g. examples in EP 0 476 014 or EP 0 703 825 and references enclosed therein). The support may be biological, non-biological, organic, inorganic, or a combination of any of these, existing as particles, strands, precipitates, gels, sheets, tubing, spheres, containers, capillaries, pads, slices, films, plates, slides, etc. The support may have any convenient shape, such as a disc, square, sphere, circle, etc. The support is preferably flat but may take on a variety of alternative surface configurations. For example, the support may contain raised or depressed regions on which the synthesis takes place. The support and its surface preferably form a rigid support on which to carry out the reactions described herein. The support and its surface is also chosen to provide appropriate light-absorbing characteristics. For instance, the substrate may be a polymerized Langmuir Blodgett film, functionalized glass, Si, Ge, GaAs, GaP, SiO_2, SiN_4, modified silicon, or any one of a wide variety of gels or polymers such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, polystyrene, polycarbonate, or combinations thereof. Other substrate materials will be readily apparent to those of skill in the art. In a preferred embodiment the substrate is flat glass or single-crystal silicon. Polynucleotide arrays (or microarrays) can be applied to expression analysis and measurement of RNA levels e.g. for a complete set of transcripts of an organism. Further, said arrays can be used for genotyping or screening of disease genes or allelic heterogeneity. Methods and materials for the production of those arrays are described in the above-mentioned literature, e.g. in EP-A-0 476 014.

In another embodiment the repeat-depleted DNA of the invention is labeled. Different procedures allow the efficient labeling of DNA probes. The labeling procedure itself is not important for a successful depletion of repetitive sequences. Such labeling methods of DNA are well-known to a person skilled in the art (e.g., Sambrook, 1989). For example, the DNA is labeled by random priming or nick translation. Further labeling techniques comprise the introduction of modified bases or nucleotides into the amplification reactions, e.g., the addition of biotinylated residues, fluorescence residues or radionucleotides to the PCR or labeling by e.g. the Biotin Chem-Link method (Boehringer, Mannheim). The labeling reaction can be affected during the

first and/or second PCR as well as between both PCR steps. Further, the DNA of the present invention can be labeled after the second PCR. A labeled repeat-depleted DNA can be used for the generation of a probe for hybridizing or for screening, e.g., in situ hybridizations, polynucleotide arrays or polynucleotide blots, e.g., DNA or RNA arrays or Southern or Northern blots.

As used herein, the term "probe" relates to polynucleotides or polynucleotide compositions, and which are used to identify another polynucleotide by hybridizing therewith. Accordingly a probe is a labeled polynucleotide which means that it can be visualized by methods well known in the art. A probe might be DNA or RNA. For example, a previously repeat-depleted DNA can be transcribed into RNA, which then can be used as probe (if appropriate). The term "probe" comprises primers, PCR products, vectors, cosmids and other polynucleotides which can be labeled. Different procedures allow the efficient labeling of probes. The labeling procedure itself is not important for a successful depletion of repetitive sequences. Such labeling methods of polynucleotides are well-known to a person skilled in the art (e.g., Sambrook, 1989). For example, DNA is labeled by random priming or nick translation. Further labeling techniques comprise the introduction of modified bases or nucleotides into an amplification reaction, e.g., the addition of biotinylated residues, fluorescence residues or radionucleotides to a PCR or primer extension or labeling can be obtained by e.g. the Biotin Chem-Link method (Boehringer, Mannheim). A labeled repeat-depleted probe can be used as a probe for hybridizing or for screening, e.g., in situ hybridizations, polynucleotide arrays or polynucleotide blots. Preferred applications and corresponding techniques are described above and in Nature genetics supplement, 21 (1999), pp. 2 - 40.

In a preferred embodiment the repeat-depleted DNA is hybridized to a second DNA. The repeat-depleted DNA can be used as a target DNA to which another DNA which has to be tested will hybridize. The tested DNA can be of every kind of DNA, e.g., genomic DNA, microdissected DNA, amplified DNA or repeat-depleted DNA etc.

Further, the DNA to be tested can also be repeat-depleted DNA of the present invention. Advantageously, the usage of repeat-depleted DNA in array technologies improves the signal to background ration after hybridization. The repeat-depleted DNA of the present invention can be used to hybridize with chromosomal DNA,

plasmids, cosmids, YACs, PACs, BACs, microdissected DNA, DNA generated by other PCR protocols, genomic DNA, or plastid DNA. Further, the repeat-depleted DNA can also be used to hybridize to further polynucleotides, e.g., RNA.

By "hybridizing" it is meant that a nucleic acid molecule hybridizes under conventional hybridization conditions, preferably under stringent conditions such as described by, e.g., Sambrook (Molecular Cloning; A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). An example of one such stringent hybridization condition in regard of a Southern blot is hybridization at 4XSSC at 65°C, followed by a washing in 0.1XSSC at 65°C for one hour. Alternatively, an exemplary stringent hybridization condition is in 50% formamide, 4XSSC at 42°C. Examples of non-stringent hybridization conditions are 4XSSC at 50°C or hybridization with 30-40% formamide at 42°C. Corresponding washing steps are washes at 37°C in 1 x SSC or at 50°C in 4 x SSC.

Thus, the present invention can also be used to generate hybridization probes for mapping the naturally occurring genomic sequence. The sequence may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. These include in situ hybridization to chromosomal spreads, flow-sorted chromosomal preparations, or artificial chromosome constructions such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price (Blood Rev. 7 (1993), 127-134) and Trask (Trends Genet. 7 (1991), 149-154). The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma, (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY. Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f) and Meinke, Science 282 (1998), 662-682. Correlation between the location of the gene on a physical chromosomal map and a specific feature, e.g., a disease, may help delimit the region of DNA associated with this feature. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. For example an sequence tagged site based map of the human genome was recently published by the Whitehead-MIT Center for Genomic Research (Hudson, Science 270 (1995), 1945-1954) on a map of the plant genome by way of the Arabidopsis genome is available from http://genome.wwz.Stanford.edu/cgi-bin/AtDB/nph-blast2atdb. Often the placement of a gene on the chromosome of another species may reveal associated marker even if the number or arm of a particular chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for interacting genes or gene defects using positional cloning or other gene discovery techniques. Once such gene or defect has been crudely localized by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

In one embodiment, the present invention relates to a repeat-depleted DNA produced by the method of the invention. The repeat-depleted DNA produced according to the present invention can be advantageously used in several applications. Once generated said DNA can be used multiple times because it can be re-amplified by stringent PCR conditions without visible use of complexity.

In another embodiment, the repeat-depleted DNA is derived from performing the method of the invention with a non-total genomic DNA, e.g. a YAC, PAC, BAC, cosmid, plasmid, or vector and therefore, the produced repeat-depleted DNA is chromosome specific DNA.

In another embodiment, the repeat-depleted DNA is ligated into vectors, plasmids, YACs, cosmids, PACs, or BACs to generate a repeat-depleted DNA library.
Said DNA library comprises a reduced ratio of competitive sequences and might have an improved background to signal ratio if it is screened. Furthermore,

advantageously the ratio of clones containing single copy sequences is increased and, therefore, the relative number of clones encoding expressed sequences in such a library is increased. Thus, a DNA library of repeat-depleted DNA will fascilitate and fasten the screening of gene or sequence.

In another embodiment, the present invention relates to a method for the production of a DNA array or DNA blot comprising immobilizing the repeat-depleted DNA of the present invention on a suitable support. Corresponding suitable supports are described above.

In another embodiment, the present invention relates to a DNA array or DNA blot, wherein the immobilized DNA is the repeat-depleted DNA of the present invention.

Accordingly, in one embodiment the DNA produced by the method of the present invention can be used for the preparation of a probe. Preferably, the probe will be labeled as described above. Such a probe can then be used for hybridizing to DNA blots or arrays as well as for chromosome painting or in situ hybridization experiments, i.e. fluoresence in situ hybridization (FISH).

According to one preferred embodiment, the present invention relates to a method for the production of a DNA library, e.g. comprising the ligation of the repeat-depleted DNA in a suitable vector, YAC, cosmid, PAC, or BAC. Methods to produce such libraries are well-known to a person skilled in the art (e.g., Sambrook, 1989).

Further, in one embodiment the invention relates to a composition comprising the repeat-depleted DNA of the invention.
The term "composition", as used in accordance with the present invention, comprises at least the repeat-depleted DNA of the invention and, optionally, further molecules, either alone or in combination, like, e.g., molecules which are capable of supporting hybridization or ligation. The composition may be in solid or liquid form or in form of (a) powder(s).

In one embodiment the composition of the present invention is a diagnostic composition comprising at least one of the aforementioned DNA, vectors, library, arrays or blots etc. and, optionally, suitable means for detection. Said diagnostic compositions may be used for methods for determining expression of DNA by detecting the presence of the corresponding mRNA which comprises isolation of mRNA form a cell and contacting the mRNA so obtained with a probe comprising a nucleic acid probe as described above under hybridizing conditions and detecting the presence of mRNA hybridized to the probe. Further, said diagnostic compositions may be used for methods for determining differences in the chromosomal composition due to translocation, inversion, etc. among normal, carrier or affected individuals. Further methods of detecting the presence of a polynucleotide according to the present invention comprises hybridization techniques are well known in the art. Accordingly, the composition can be used for pre- and postnatal diagnostic.
Advantageously, the diagnostic composition is used for the identification of a polymorphism, a structural and complex chromosomal rearrangement, for pre- or postnatal diagnostic, the identification of supernumerary marker chromosomes (SMC), or visualization of arcocentric chromosomes, e.g. of the p-arms of arcocentric chromosomes, tumor cytogenetics, etc.
The composition of the present invention can be provided in parts or together, and each part, e.g. a vector or a template DNA, can be packaged individually.

In one embodiment the present invention relates to a kit comprising the repeat-depleted DNA of the invention and/or instructions for performing the method of the present invention. Advantageously, the kit of the present invention further comprises, optionally, DNA polymerases (such as Taq polymerase, Pwo polymerase, and/or ThermoSequenase), a repetitve template DNA as well as (a) reaction buffer(s) and/or storage solution(s). Furthermore, parts of the kit of the invention can be packaged individually in vials or in combination in containers or multicontainer units. The kit of the present invention may be advantageously used for carrying out the method of the invention and could be, inter alia, employed in a variety of applications referred to above, e.g. in diagnostic kits or as research tools. Additionally, the kit of the invention may contain means for detection suitable for scientific and/or diagnostic purposes. The manufacture of the kits follows preferably standard procedures which are known

to the person skilled in the art. In one embodiment a PCR kit is used for the method of the present invention. Such a PCR kit can comprise, e.g., primers, a DNA-DNA polymerase (e.g., Taq polymerase, ThermoSequenase or Pwo polymerase), deoxyribonucleotides, $MgCl_2$, $MnCl_2$ and/or repetitive sequences comprising template DNA.

Several documents are cited throughout the text of the specification. Each of the documents cited herein including any manufacturer's specifications, instructions, etc. are hereby incorporated by reference; however, there is no admission that any document cited is indeed prior art of the present invention.

Throughout the specification and the claims which follow, unless the context requires otherwise, the word "comprise" and variations such as "comprising" will be understood to imply the inclusion of a stated integer or a step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The figures show:

Figure 1 shows the hybridization patterns of total genomic DNA. In **(A)** the hybridization result of untreated genomic DNA hybridized without Cot-1 DNA and without preannealing is shown. This figure shows total genomic DNA labeled with standard nick translation and hybridized without preannealing and addition of Cot-1 DNA. As expected the large heterochromatic blocks show a very intense staining. **(B)** The same DNA labeled by said PCR, hybridized under the same conditions as mentioned above: The heterochromatic blocks are already not stained any more, however, the chromosome arms show no homogenous painting pattern but R-banding similar banding pattern. **(C)** The DNA shown in (A) after a second round of said PCR amplification. **(D)** The DNA shown in (B) after a third round of said PCR. The chromosome arms show a homogeneous staining pattern.

Figure 2 shows the hybridization pattern of a chromosome 14 suppressed by three rounds of said PCR. Hybridization pattern of chromosome 14 after depletion of repetitive sequences. **(A)** shows the DAPI image, **(B)** the respective Cy3.5 image,

and **(C)** a "composite" (overlay of DAPI and Cy3.5 image). After depletion of repetitive sequences, specific hybridization to chromosome 14 can be achieved without Cot-1 DNA and without preannealing. Note that the centromeres and the p-arms of other acrocentric chromosomes do not show hybridization signals.

Figure 3 shows that with the YAC highly specific signals can be achieved. Shown is a YAC clone after several rounds of said PCR.

The invention will now be described by reference to the following biological examples which are merely illustrative and are not to be provided as a limitation of scope of the present invention.

Example 1: Generation of a repeat-depleted DNA probe

A number of DNA probes can be processed by the method described, including chromosome-specific painting probes generated by microdissection or flow-sorting, any region-specific microdissected probe, cosmids, YACs, PACs, BACs, etc.

The CEPH-YAC 933a5 used in this study was provided by Dr. Thomas Haaf from the Max-Planck Institute of Human Genetics in Berlin (http://www.mpimg-berlin-dahlem.mpg.de). The YAC was placed on a pulsed-field gel, after running the human insert was cut from the gel and amplified via said PCR.

In addition, this approach should also work for probes for other species, such as mouse-probes. The DNA is amplified using the method of the present invention. The reaction mixture contained said DNA (usually about 100 ng, however, amount may vary), 5 µl 10x-PCR-buffer (without $MgCl_2$), 4 µl 25mM $MgCl_2$, 2 µl 5mM dNTP, 5 µl 6MW-primer 17.0µM (final concentration: 1.7 µM), 0.5 µl Taq Polymerase (2.5 units) and add 50 µl with add H_2O. PCR with 6MW Primer:[5´-CCG ACT CGA GNN NNN NAT GTG G-3´] [SEQ ID NO: 1] was then performed by the following conditions: 5 min at 93°C, followed by five cycles of 1 min at 94°C, 1.5 min at 30°C, 3 min transition 30°C-72°C (this refers to the time ramp mentioned above), and 3 min extension at 72°C, followed by 35 cycles of 1 min at 94°C, 1 min at 62°C, and 3 min at 72°C, with an addition of 1 sec/cycle to the extension step and a final extension of 10 min.

However, the conditions of the PCR can be varied as following: The 30°C-72°C transition is elongated to a time longer than 3 min, e.g. 6 min transition time. Cycling conditions are: (1) 94°C for 5 min; (2) five low-stringency cycles of 94°C for 1 min, 30°C for 1.5 min, 6 min transition 30°C-72°C, and 72°C for 3 min; (3) 35 high-stringency cycles of 94°C for 1 min, 62°C for 1 min, and 72°C for 3 min, with the addition of 1 s per cycle to the extension time; and (4) a final extension of 5 min at 72°C.

Further, the number of low-stringency cycles can be varied. Following cycling conditions can be performed: (1) 94°C for 5 min; (2) five to ten low-stringency cycles of 94°C for 1 min, 30°C for 1.5 min, 6 min transition 30°C-72°C, and 72°C for 3 min; (3) 35 high-stringency cycles of 94°C for 1 min, 62°C for 1 min, and 72°C for 3 min, with the addition of 1 s per cycle to the extension time; and (4) a final extension of 5 min at 72°C. Additionally, the 3 min transition 30°C-72°C can be omitted.

Furthermore, the PCR can be done with different primer pairs, e.g. DNA-templates can be amplified with a more stringent version of said PCR called CTA4-PCR (Craig et al. 1997). About 100 ng of the PCR products were amplified by the CTA4-PCR technique according to published protocols (Craig et al. 1997). The conditions were essentially the same as for said PCR. The only differences include the replacement of primer 6MW by the primer 5'-CTA CTA CTA CTA CCG ACT CGA G-3' [SEQ ID NO: 2], annealing was performed at 53°C with or without a time extension, and the number of normal cycles was 36.

Example 2: Probe labeling

Different procedures allow the efficient labelling of DNA-probes. The labeling procedure itself is not important for a successful depletion of repetitive probes.

Labeling of individual painting probes and the above-mentioned probe pools was carried out by CTA4-PCR. In biotin- or digoxigenin- (both Boehringer Mannheim) and Cy3- (Amersham Pharmacia Biotech) labeling reactions, 0.08 mM dTTP and 0.02 mM of the respective fluorochrome-dUTPs were used. Cy5 was labeled with 0.068 mM dTTP and 0.032 mM Cy5-dUTP (Amersham Pharmacia Biotech). The concentration of dATP, dGTP, and dCTP was 0.1 mM each. FITC labeling was

achieved with 0.032 mM FluorX-dCTP (Amersham Pharmacia Biotech), 0.068 mM cCTP, and 0.1 mM each of dATP, dGTP, and dTTP.

Any variation of aforementioned PCR may be used for the probe labeling. The exact probe and fluor concentrations were described by Eils (1998). Standard nick-translation, random-primed labeling, or Chemical labeling is also suitable.

Example 3: Fluorescence in situ hybridization with repeat-depleted DNA probes

A series of FISH experiments was done with various probe concentrations. Probe concentrations in experiments with a single painting probe or more probes was in the range of 10-20 ng DNA per probe per microliter of hybridization solution.

Five micrograms of salmon testis DNA (Sigma) was added to all probe mixtures. Unlabeled Cot-1 DNA was added only in control FISH experiments of untreated probes hybridized using standard protocols. The probe mixtures were precipitated with ethanol and resuspended in a hybridization solution containing 50% formamide, 2xSSC, and 10%-20% dextran sulfate. Probe mixtures containing depleted probes were denatured but not reannealed and hybridized for one or two nights at 37°C to metaphase chromosome spreads. Hybridization with untreated probes was essentially the same, except that probes were allowed to preanneal for 20 min or longer.

After hybridization, the slides were washed three times with 4xSSC, 0.2% Tween 20 at 45°C and then three times with 1xSSC at 60°C. In experiments with biotin- or digoxigenin-labeled DNA probes, unspecific binding sites were blocked with 3% BSA in 4xSSC, 0.2% Tween for 30 min in a moist chamber. Biotin-labeled DNA probes were detected with Avidin-Cy3.5 (Amersham Pharmacia Biotech). Digoxigenin-labeled probes were detected using either one layer of anti-digoxigenin-Cy7 (Amersham Pharmacia Biotech) or with a different, two layer system consisting of anti-dig rabbit (Sigma) in the first step and anti-rabbit Cy5.5 (Amersham Pharmacia Biotech) in a second step. After final washes with 4xSSC, 0.2% Tween at 45°C, slides were counterstained with 4,6-diamidino-2-phenylindole (DAPI) and mounted in phenylenediamine antifade medium. The fluors used are not decisive for the outcome of the experiment, it should work with any fluor from any manufacturer.

The images of the hybridized metaphase spreads were captured using a Leica DMRXA-RF8 microscope equipped with a cooled Sensys CCD camera

(Photometrics) controlled by Leica QFISH software (Leica Microsystems Imaging Solutions Ltd.), as described elsewhere (Eils et al., 1998).
Further, the images were analyzed using the Leica image analysis software programs QFISH, QWIN, and MCK (Leica Microsystems Imaging Solutions Ltd.).

Regions rich with repetitive DNA (e.g. the centromeres of most chromosomes) show an extensive staining (see Fig. 1). After the first round of PCR, the centromeres are not stained any more, but show a "hybridization gap". This indicates a sufficient suppression of alphoid sequences. However, not all repetitive sequences are sufficiently suppressed as seen by the R-banding like hybridization pattern on most chromosome arms. The hybridization pattern improves after subsequent rounds of PCR and yields satisfactory results after the third round, and the "usual-standard" result after the fourth round of PCR.

Chromosome 14 presents as all acrocentric chromosomes a special case because the short arm contains sequences coding for ribosomal RNA. Thus, this chromosome is particularly rich with repetitive sequences. As Fig. 2 shows, the PCR was capable of generating a highly specific probe.
In order to test whether the procedure would also work for smaller probes, CEPH-YAC 933a5 was reamplified three times and hybridized to metaphase spreads. Fig. 3 shows that with the YAC highly specific signals can be achieved.

The results shown in Fig. 1-3 were all obtained without the addition of Cot-1 DNA and without preannealing. Furthermore, images shown are the source images without image processing, thus they really reflect a highly specific hybridization pattern.

The assessment of the quality of individual repetitive sequence–free painting probes showed that a successful depletion could be fulfilled and specific hybridization to the target chromosome if hybridized without Cot-1 DNA and without preannealing were obtained. Further, the signal-to-background ratio fall at least within the range of the one obtained with the untreated painting probe hybridized under standard conditions. These results could be obtained for a large number of different probes, as total genomic DNA, chromosome 14 painting probe, or CEPH-YAC 933α5.

References

Carter NP, Bioimaging 4 (1996), 41-51.

Chen-Liu LW, Huang BC, Scalzi JM, Hall BK, Sims KR, Davis LM, Siebert PD, Hozier JC, Genomics 30 (1995), 388-390.

Cook HJ, Hindley J, Nucl Acids Res 6 (1979), 3177-3197.

Craig JM, Kraus J, Cremer T, Hum Genet 100 (1997), 472-476.

Dunham I, Lengauer C, Cremer T, Featherstone T, Hum Genet 88 (1992), 457-462 (1992).

Eils R, Dietzel S, Bertin E, Schröck E, Speicher MR, Ried T, Robert-Nicoud M, Cremer C, Cremer T, J Cell Biol 135 (1996), 1427-1440.

Eils R, Uhrig S, Saracoglu K, Sätzler K, Bolzer A, Chassery JM, Petersen I, Ganser M, Speicher MR, Cytogenet Cell Genet 82 (1998), 160-171.

Guan XY, Meltzer PS, Trent JM, Genomics 22 (1994), 101-107.

Kallioniemi A, Kallioniemi O-P, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D, Science 258 (1992),818-821.

Landegent JE, Jansen in de Wal N, Dirks RW, Baas F, van der Ploeg M, Hum Genet 77 (1987), 366-370.

Lengauer C, Riethman HC, Speicher MR, Taniwaki M, Konecki D, Green ED, Becher R, Olson MV, Cremer T, Cancer Res 52 (1992), 2590-2596.

Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC, Hum Genet 80 (1988), 224-234.

Lichter P: Multicolor FISHing, Trends Genet 13 (1997), 475-479.

Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual, 2nd Ed (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).

du Manoir S, Speicher MR, Joos S, Schröck E, Popp S, Döhner H, Kovacs G, Robert-Nicoud M, Lichter P, Cremer T, Hum Genet 90 (1993), 590-610.

Müller S, O'Brien P, Ferguson-Smith MA, Wienberg J, Hum Genet 101 (1997), 149-153.

Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J, Gray JW, Proc natl Acad Sci, USA 85 (1988), 9138-9142.

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG, Nature Genet 20 (1998), 207-211.

Rouquier S, Trask BJ, Taviaux S, van den Engh G, Diriong S, Lennon GG, Giorgi D, Nucl Acids Res 23 (1995), 4415-4420.

Schröck E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, Garini Y, Ried T, Science 273 (1996), 494-497.

Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Brenner A, Döhner H, Cremer T, Lichter P, Genes Chrom Cancer 20 (1997), 399-407.

Speicher MR, Ballard SG, Ward DC, Nature Genet 12 (1996), 368-375 (1996).

Szabo P: In situ hybridization with radiolabelled probes, in Verma R (ed): Advances in Genome Research: Unfolding the Genome, vol 1. Greenwich, CT, JAI Press, 1992, pp 175-216

Telenius H, Pelmear AH, Tunnacliffe A, Carter NP, Behmel A, Ferguson-Smith MA, Nordenkj¨ld M, Pfragner R, Ponder BAJ, Genes Chrom Cancer 4 (1992), 257-263.

Wienberg J, Adamski E, Yang F, Müller S, Ferguson-Smith MA: Chromosome painting without competitor DNA. Trends Genet Tech Tips Online (http://www.elsevier.nl/locate/tto) (1997).

Claims

1. A method for the generation of repeat-depleted DNA comprising amplifying repetitive template DNA by a first polymerase chain reaction (PCR), wherein the hybridization step is a low stringency hybridization step and a second PCR following the first PCR, wherein the hybridization step of said second PCR is a high stringency hybridization step.

2. The method of claim 1 comprising a preannealing step.

3. The method of claim 1 or 2 wherein the transition time between the hybridization step and the elongation step of said first PCR is between about 2 minutes and about 8 minutes.

4. The method of any one of claims 1 to 3 wherein the transition time between the hybridization step and the elongation step of said first PCR is between about 3 minutes and about 6 minutes.

5. The method of any one of claims 1 to 4 wherein said first PCR comprises more than 3 amplification cycles.

6. The method of claim 5 wherein said first PCR comprises between 5 and 10 cycles.

7. The method of any one of claims 1 to 6 wherein the time period of the hybridisation step of the first PCR is between about 1 and 8 minutes.

8. The method of any one of claims 1 to 7, wherein said first and/or second PCR is/are repeated at least once.

9. The method of any one of claims 1 to 8, wherein said repetitive template DNA is derived from higher or lower eukaryotes.

10. The method of claim 9, wherein said repetitive template DNA is derived from mammals or plants.

11. The method of any one of claims 1 to 10, wherein said repetitive template DNA is a YAC, cosmid, PAC, BAC, plasmid, microdissected DNA, DNA isolated from hybrid cell lines, DNA generated by other PCR protocols, plastid DNA or is total genomic DNA.

12. The method of any one of claims 1 to 11 further comprising the step of ligating said repeat-depleted DNA into vectors, YACs, PACs, BACs, or plasmids.

13. The method of any one of claims 1 to 12, comprising the step of immobilizing the repeat-depleted DNA on a suitable support.

14. The method of claim 13 wherein the support is a DNA array or a DNA blot.

15. The method of any one of claims 1 to 14 further comprising labeling said repeat depleted DNA.

16. The method of claim 15, wherein said labeling is effected after the first and/or second PCR.

17. The method of claim 15 or 16 wherein the labeled repeat-depleted DNA is hybridized to a second DNA.

18. The method of claim 17 wherein said second DNA is chromosomal DNA, total genomic DNA a cosmid, plasmid, YAC, PAC, BAC, microdissected DNA, DNA isolated from hybrid cell lines, DNA generated by other PCR protocols or plastid DNA.

19. Use of the method of any one of claims 1 to 18 for the generation of a repeat depleted DNA.

20. A repeat depleted DNA produced by the method of any one of claims 1 to 18.

21. The repeat-depleted DNA of claim 20, wherein the repeat-depleted DNA is a chromosome specific DNA.

22. A method for the generation of a DNA library comprising ligating the repeat-depleted DNA of claim 20 or 21 into vectors, YACs, PACs, BACs, cosmids or plasmids.

23. A method for the production of a DNA array or DNA blot comprising immobilizing the repeat-depleted DNA of claim 20 or 21 on a suitable support.

24. A DNA array or DNA blot, wherein the immobilized DNA is the repeat-depleted DNA of claim 20 or 21 or produced according to the method of claim 23.

25. A DNA library comprising the repeat-depleted DNA of claim 20 or 21 or generated according to the method of claim 22.

26. Use of the repeat-depleted DNA of claim 20 or 21 for the preparation of a probe.

27. The use of claim 22 wherein the probe is a probe for in situ hybridization, for chromosome specific paiting, or for hybridizing of DNA arrays or DNA blots.

28. Use of the repeat-depleted DNA of claim 20 or 21 for the construction of a DNA library.

29. Use of the repeat-depleted DNA of claim 20 or 21 for hybridizing with chromosomal DNA, cosmids, YACs, PACs, BACs, plasmids, microdissected DNA, DNA isolated from hybrid cell lines, DNA generated by other PCR protocols, plastid DNA or total genomic DNA.

30. The method of claim 17 or the use of claim 29, wherein said hybridizing is an in situ hybridization.

31. A composition comprising the repeat-depleted DNA of claim 20 or 21.

32. The composition of claim 31 , which is a diagnostic composition.

33. A kit comprising the repeat-depleted chromosomal DNA of claim 20 or 21, the DNA library of claim 25 and/or the DNA array or the DNA blot of claim 24.

34. A kit comprising instructions for performing the method of any one of claims 1 to 18 or the method of any one of claims 22 or 23.

35. The kit of claim 34 wherein said kit comprises further oligonucleotides and, optionally, a DNA-DNA polymerase.

36. Use of a PCR-kit for the method of any one of claims 1 to 18.

37. Use of the kit of claim 34 or 35 for the method of any one of claims 1 to 18 or for the generation of DNA libraries, for in situ hybridization, for the production of DNA arrays or DNA blots.

38. The method of claim 30 or the use of any one of claims 27, 30 or 37, wherein said in situ hybridization is a fluorescence in situ hybridization (FISH).

WO 01/29252 PCT/EP00/10268

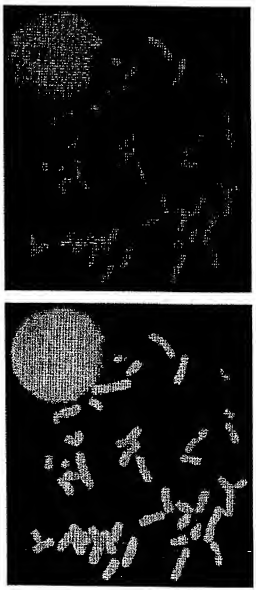


Fig. 1a

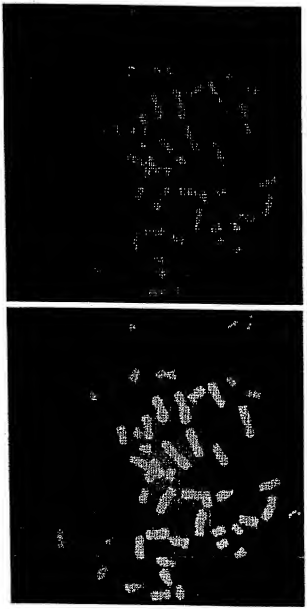


Fig. 1b

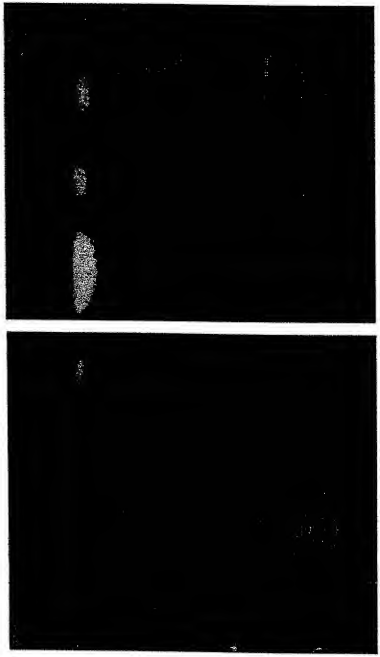


Fig. 1c

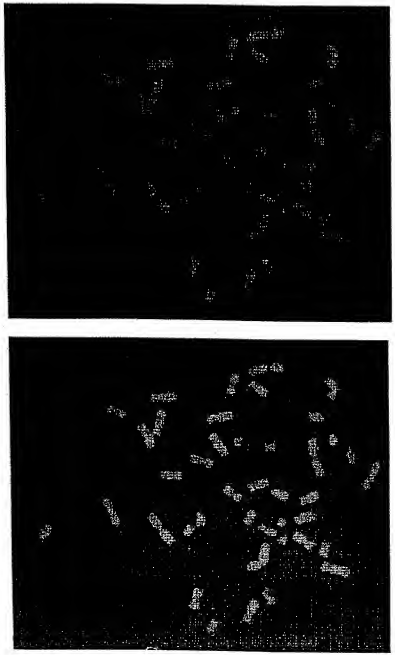
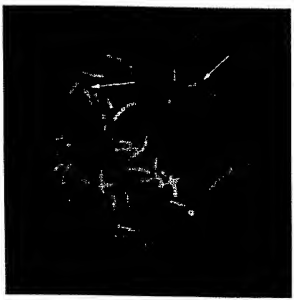


Fig. 1d

WO 01/29252 PCT/EP00/10268

Fig. 2

Fig. 3



WO 01/29252 PCT/EP00/10268

SEQUENCE LISTING

<110> SPEICHER, Michael

<120> A method for the generation of repeat-depleted DNA

<130> D2045PCT

<140>
<141>

<160> 2

<170> PatentIn Ver. 2.1

<210> 1
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial sequence

<400> 1
ccgactcgag nnnnnnatgt gg 22

n=a, c, g or t

<210> 2
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial sequence

<400> 2
ctactactac taccgactcg ag 22